



**Eur päisch s  
Patentamt**

**Eur p an  
Patent Office**

**Office européen  
des brevets**

**Bescheinigung**

**Certificate**

**Attestation**

Die angehefteten Unterla-  
gen stimmen mit der  
ursprünglich eingereichten  
Fassung der auf dem näch-  
sten Blatt bezeichneten  
europäischen Patentanmel-  
dung überein.

The attached documents  
are exact copies of the  
European patent application  
described on the following  
page, as originally filed.

Les documents fixés à  
cette attestation sont  
conformes à la version  
initialement déposée de  
la demande de brevet  
européen spécifiée à la  
page suivante.

**Patentanmeldung Nr. Patent application No. Demande de brevet n°**

**02405951.1**

Der Präsident des Europäischen Patentamts;  
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets  
p.o.

**R C van Dijk**





Anmeldung Nr:  
Application no.: 02405951.1  
Demande no:

Anmeldetag:  
Date of filing: 06.11.02  
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

INVENTIO AG  
Seestrasse 55  
CH-6052 Hergiswil NW  
SUISSE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:  
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.  
If no title is shown please refer to the description.  
Si aucun titre n'est indiqué se referer à la description.)

Video camera monitoring of escalators and moving walks

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s)  
revendiquée(s)

Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/  
Classification internationale des brevets:

B66B29/02

Am Anmeldetag benannte Vertragsstaaten/Contracting states designated at date of  
filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LI LU MC NL PT SE SK TR



**Video camera monitoring of escalators and moving walks.****Description**

- 5 The present invention **refers** to video camera monitoring of escalators and/or moving walks according to the definition of the independent claims.

Such monitoring systems are well known in different  
10 embodiments as escalator start lock or escalator restart control. Through such monitoring systems the escalator restart after a voluntary or erroneous actuation of the emergency stop or other safety devices must remain blocked, until no person or object are present in the monitored  
15 field of the safety device.

In particular, the escalator restart control conditions are established by the norm E115: an escalator, which is used to transport people in environments such as railway  
20 stations, shopping centres etc., should be monitored for security reasons. The monitoring is restricted to the case, where an escalator is stopped, and a safe restart is required. A safe restart may only be performed in situations where the escalator is repeatedly tested for  
25 emptiness, i.e. there are no persons or obstacles on the moving part and the entry regions of the escalator. The required period of emptiness is typically adjusted to 10 seconds. Over this period repeated checks for emptiness may be performed every 0.1 seconds.

30

Monitoring systems of this type are for example described in EP 1013599, JP 10236757 and in JP 10265163.

**EP 1013599** discloses a monitoring system for escalator restart control, which detects the presence of persons or objects on the escalator through a set of cameras situated above the escalator. Practical experiments have  
5 demonstrated that this system does not work in the case of strong sun irradiation, faint diffused light and in the case of rain, drizzle and fog, and that under these circumstances an unambiguous perception of the emptiness of the escalator cannot be assured.

10

**JP 10236757** shows a remote supervisory system whereby without dispatching a clerk in charge to a setting site of an escalator, a moving guide device is remotely supervised, start and stop control can be performed. This remote  
15 supervisory system comprises ITV cameras supervising an escalator and its periphery and a central controller provided in a remote area to control a start/stop of the escalator.

20 **JP 10236757** shows an escalator controller to judge a phenomenon and to speedily correspond to a falling accident, etc., in accordance with a picked-up picture image of an escalator and its periphery.

25 Both **JP 10236757** and **JP 10236757** disclose monitoring systems, which do not work properly under special illumination conditions and cannot guarantee the unequivocal perception of the emptiness of the escalator. In particular, shadows or dirty spots on the escalator can  
30 be eventually confused with people or objects lying on the escalator itself.

The **object of the present invention** is to conceive a monitoring of obstacles and persons on escalators and/or

moving walks, which allows a reliable and univocal detection of persons or obstacles lying in the monitored field of the escalator and/or moving walk.

- 5 According to the present invention **this object is achieved** by a monitoring system for the detection of obstacles and persons on escalators and/or moving walks comprising at least one video camera for the acquisition of stereoscopic images.

10

With the word **stereoscopic images** is meant a pair of pictures of the same field of view taken by two cameras situated at slightly different positions or taken by the same camera placed in two slightly different positions, so  
15 that the same field of view is imaged under two slightly different angles. The objects on the escalator, which should be detected, have the property of being closer to the camera than the escalator on which they are placed. The advantage of stereoscopic images is that these objects  
20 appear at different positions in the pair of stereoscopic images. Disturbances like dirt or inscriptions on the escalator appear on the same position in the pair of stereoscopic images, so that it is univocally possible to detect the presence of objects and persons on the  
25 escalator.

As **obstacles or persons** are understood objects and bodies, whose dimensions are such as to endanger the safe operation of the escalator and/or moving walk.

30

In preferred embodiments of the invention pairs of video cameras are located above the escalator or in the escalator balustrade. This embodiment exhibits the advantage that an optimal field of view of the escalator can be achieved,

i.e. under a view angle of  $45^\circ$  the obstacles and persons are neither too close to the camera (too big in the image) nor too far away (too small in the image).

If the escalator is very long, more than one pair of  
5 cameras is necessary to monitor conveniently the whole stroke of the escalator.

In another preferred embodiment the monitoring system includes a processing unit to process the stereoscopic  
10 images. This embodiment exhibits the advantage that the monitoring system can automatically process the acquired images and can autonomously come to the decision whether obstacles are present on the escalator or not.

15 In this context with **processing** the stereoscopic images is meant any operation preferably performed on digital images such as loading, storing, comparison, differencing, rectifying, warping, reconstruction, segmentation, grouping, edge detection, Hough transformations, extraction  
20 etc. as described below in the detailed description of the invention.

The processing unit can be a personal computer or a standardised non-expensive processor integrated in the  
25 camera or in any other part of the escalator equipment needing no special device to be mounted.

In another preferred embodiment the processing unit and the cameras can be connected together by linking means and even  
30 with the escalator controller. This embodiment exhibits the advantage that the monitoring system can automatically process the acquired images, can autonomously come to the decision whether obstacles are present on the escalator or not and can finally restart automatically the escalator



basing on the obtained information. With **linking means** is understood any physical means, such cables, signals or a data exchange bus, which allows data to be exchanged and transmitted between two or more acquisition, processing and  
5 controlling units.

According to the present invention **the object is also achieved by a method** for the detection of obstacles and/or persons on escalators and/or moving walks, whereby at least  
10 one video camera acquires stereoscopic images and a processing unit processes these images. The advantage of this method is that it is easy to perform and reliable.

Another preferred embodiment of the invention refers to a  
15 **computer program product** for the detection of obstacles and/or persons on escalators and/or moving walks, which loads in a processor and processes stereoscopic images of the escalator and/or moving walk. The advantage of the computer program product is that it is loadable everywhere,  
20 locally or remotely in a central server and that an update is easy to perform.

Preferred embodiments of the invention are described in detail below with reference to the following schematic  
25 drawings:

Fig.1: A complete representation of the escalator equipped with the monitoring system according to the invention.

30 Fig. 2: Monitoring system whereby a pair of cameras is placed in the escalator balustrade.

Fig. 3: Monitoring system whereby a pair of cameras is mounted at the top of two posts placed along the escalator.

Fig. 4: Collaboration diagram for image data exchange using a shared memory.

5 Fig. 5: Data flow diagram for the full system.

**Fig. 1** shows a complete representation of the escalator equipped with the monitoring system according to the invention. On the escalator 1 is standing a person 2, which is in the field of view of a pair of video cameras 3.1 and 3.2 placed at slightly different positions above the escalator. The cameras can therefore acquire pairs of stereoscopic images of the escalator.

15 Image acquisition is performed using pairs of cameras, where the number  $n_c$  of cameras required depends on the height of the staircase  $H$ . An estimate is given by  $n_c = 4 + H$ , where  $H$  is the height of the escalator. For example, for  
20 a staircase spanning four meters in height four stereo camera pairs, i.e. eight cameras at all, are necessary.

For the internal part of the escalator two cameras with a focal length of 6 millimetres (nominal) in turned mode,  
25 i.e. the vertical image extension is larger than the horizontal one, are suggested. The entry regions on the top and on the bottom of the escalator also require cameras with a focal length of 6 millimetres (also in turned mode).

30 Environment and escalator parameters influencing the placement and the number of the cameras are for example the length of the escalator, which can be up to 100 meters. The escalator may be located in- or outdoors with or without covering. Escalator stairs may be coloured or hold

inscriptions. An opaque object of cylindrical shape and minimum size of 0.15 meters in diameter and 0.15 meters in height must be detected as a necessary requirement. The illumination may vary over the escalator area, a minimum  
5 illumination is given as 50 Lux for indoor placement and 15 Lux for outdoor placement.

**Fig. 2** shows a preferred embodiment of the monitoring system whereby a pair of cameras is placed in the escalator  
10 balustrade.

**Fig. 3** shows a preferred embodiment of the monitoring system whereby a pair of cameras is mounted at the top of two posts placed along the escalator.

15

B/W **cameras** and DFG/BW1 frame grabbers manufactured by The Imaging Source with a progressive scan CCD image sensor can be used.

An important additional requirement are lenses controllable by voltage. This requirement stems from the highly varying  
20 illumination conditions.

The used lens is a Cosmincar lens, type H612ER, with a focal length  $f$  of 6 millimeters. The aperture opening is controllable from  $f/1.2$  to  $f/360$  through variation of the  
25 control voltage in the range 1.5 to 5 volts. The aperture is controlled using the NuDAQ 6208 multi-channel analogue output card.

The cameras are connected through the linking means 4 (for  
30 example Hirose cables) to a **processing unit** 5, which processes the digitalized stereoscopic images taken by the video camera pair. Thanks to algorithms described below, the processing unit detects the presence or not of the person on the escalator. Detection is based on differencing

rectified stereo pair images, where a warping transform is to overlay the left onto the right image, and vice versa. The 3D camera positions are obtained through model based pose estimation and disparity is used to obtain the warping  
5 transform.

In particular, the task is to detect objects on an escalator, which can be considered as a moving background, under real-world illumination conditions. The suggested  
10 solution consists of model based background reconstruction, perspective warping of one image to the other in a stereo setup, and the final detection of differences in an image pyramid. Specifically, we employ a model based staircase pose estimator based on grouping of line features by the  
15 use of geometric invariants. The detection is based on measuring absolute pixel differences between unwarped and warped images. Image differences are represented in an image pyramid according to Peter J. Burt, Tsai-Hong Hong, and Azriel Rosenfeld, "Segmentation and Estimation of Image  
20 Region Properties Through Cooperative Hierarchical Computation", IEEE Transactions on Systems, Man and Cybernetics, 11(12):802-809, December 1981, and segmented into background (staircase) and foreground (obstacles) employing the algorithm suggested in M. Spann and R.  
25 Wilson, "A Quad-Tree Approach to Image Segmentation which Combines Statistical and Spatial Information", Pattern Recognition, 18(3/4):257-269, 1985.

Image processing is performed on **PC-class machines** (Intel  
30 Pentium). The number of PC boxes can be higher than one. In a preferred embodiment each PC box is responsible for two stereo pairs, i.e. is connected with four cameras. Each PC is equipped with one NuDAQ 6208 and two DFG/BW1 cards.

The software is written in C++ and runs under the Linux operating system. Efficient image, computer vision and matrix algebra algorithms are provided by the used Intel Performance Primitives Library.

5

The main software components are:

- **Acquisition**, including aperture control.
- **Calibration** of camera and system (offline).
- **Monitoring**, state estimation and detection (online).

10 The detection part is the time-critical part performed at escalator service time (online), whereas the calibration part is done beforehand, i.e. at escalator assembly time (offline).

15 **Acquisition** performs two tasks:

- Providing grabbed images at some negotiated shared memory.
- Control of the aperture based on image properties, e.g. maximization of the information content in the staircase region of interest (ROI).

20

**Fig. 4** shows a collaboration diagram, which explains the communication between acquisition component and processes requiring images, i.e. the offline and online components.

25 The basic principles for **synchronisation and communication** are:

- Components are Unix processes.
- Image data are exchanged using shared memories.
- Synchronisation for shared memory access uses semaphores.

30

Signalling between processes uses message queues.

For the **calibration and monitoring** part, there are seven main tasks identified:

- Radial/tangential undistortion.
- Motion segmentation and ROI identification.
- 5 • Edge and line extraction.
- Geometric matching, model/data line correspondence, pose estimation.
- Disparity calculation, warping table setup.
- Staircase state estimation.
- 10 • Image warping, segmentation, connected components labelling, decision support.

The undistortion task is required in the offline and online parts. The next first four components can be summarized as offline component, whereas the last two are the online  
15 component.

**Fig. 5** shows the data flow for the system, with the above mentioned components. The external data stores provide undistortion parameters from internal calibration and a CAD  
20 model of the staircase, i.e. a list of points and lines. Output, which is the result of detection, goes to another external data store. The main components: acquisition, offline and online are grouped in shaded areas. Undistortion is applied to images gathered by both, the  
25 online and the offline component.

As stated above, the main components of the presented system are the acquisition part, the offline (or calibration) part and the online (or detection) part. The  
30 most interesting subparts, i.e. **geometric matching** (establishing correspondences between 2D-Data and 3D-Model) in the offline part and **detection from stereo images** in the

online part, will be now discussed in some detail in the following.

In model-based pose estimation parameters describing  
5 relative orientation and position, i.e. the extrinsic  
camera parameters, are found using **correspondence between  
data and model**. In our case, the data are 2D lines  
extracted from single images and the model is a 3D  
10 wireframe object. Nearly horizontal lines are derived from  
the image data using standard edge detection based on  
directional image gradient and Hough transform techniques.  
To establish correspondence between data and model lines  
for each image in the stereo pair, and furthermore, between  
the two stereo pairs, the following matching procedure is  
15 applied (**grouping based on cross ratio**).

The first step in matching is to identify possible  
correspondences between data and model lines. Under  
perspective projection, ratios of ratios of lines and  
20 ratios of ratios of angles, the so called cross ratios, are  
invariant. We employ cross ratios to identify groups of  
four lines out of a larger set of possible lines. Such a  
group of 4 lines, which in our case is characterized by the  
cross ratio obtained for the intersection points with an  
25 approximately orthogonal line, serves as a matching  
candidate to the staircase pattern. The definition for the  
cross ratio for four points  $p_1, \dots, p_4$  on a line is given as:

$$\text{Cr}(p_1, \dots, p_4) = [(x_3 - x_1)(x_4 - x_2)] / [(x_3 - x_2)(x_4 - x_1)],$$

30

where the  $x_1, \dots, x_4$  are the corresponding positions of each  
point on the line.

The following strategy for selecting data lines which are good candidates for correspondence to model lines was employed:

- Calculate the theoretical cross ratio, e.g. for four  
5      equally spaced points on a line this is  $Crt = 4/3$ .
- Detect a reasonable set  $L$  (of size  $N$ ) of close to  
horizontal lines from the data.
- Calculate intersection points of those lines with a  
close to vertical line.
- 10    • Calculate all  $M = \binom{N}{4}$  four-element subsets of lines  $l_i \subset$   
 $L, i=1, \dots, M$ .
- Calculate all cross ratios  $c_i$  corresponding to sets  $l_i$ .
- Sort the  $l_i$  with respect to  $|c_i - Crt|$  (in ascending  
order).
- 15    Only a portion of the sorted groups, corresponding to those  
of lower distance to  $Crt$ , is input to the pose estimation  
step, which is described below (**estimation of position and  
orientation**).
- 20    Corresponding groups of lines are input to a procedure  
similar to RANSAC as described in M.A. Fischler and R.C.  
Bolles, "Random Sample Con-census: A Paradigm for Model  
Fitting with Applications to Image Analysis and Automated  
Cartography", Communications of the ACM, 24(6):381-395,  
25    1981. Grouping based on cross ratio delivers improved  
sampling for RANSAC and reduces the number of necessary  
iterations. The basic idea in RANSAC is, that RANSAC uses  
as small an initial data set as feasible and enlarges this  
set with consistent data when possible. The required number  
30    of random selections  $n_s$  of samples with a size of  $s$   
features is given by Fischler and Bolles as:



$$n_s = \log(1-p_c) / \log(1-p_i^s),$$

where  $p_c$  is the probability, that at least one sample of  $s = 4$  lines is free from outliers. The probability that any selected sample is an inlier is denoted by  $p_i$ . In our case, due to the improved sampling based on cross ratio, we can safely assume a high  $p_i$ , e.g.  $p_i=0,8$ , and choosing  $p_c=0,99$ , we obtain a number of necessary RANSAC iterations as low as  $n_s = 9$ .

Verification of the pose is based on the procedure devised by David G. Lowe, "Fitting Parameterized 3-D Models to Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, 13(5):441-450, May 1991. Lowe approaches the problem of derivation of object pose from a given set of known correspondences between 3D-model lines and 2D image lines by linearization of projection parameters and application of Newton's method. The result of the pose estimation step are two transformations from world to camera coordinate system, i.e. three translational and three rotational parameters for each camera.

The **detection from stereo images** involves detector calibration, i.e. derivation of disparity and derivation of the two-dimensional warping transform, and the detection itself, i.e. warping of one image to the other, differencing of warped and unwarped images and, finally, segmentation of the difference image in order to obtain a decision.

The **warping transform** is found from the staircase model and the two world to camera coordinate system and projective transforms obtained by the pose estimation procedure mentioned above. A perspective warping transform provides us with two warping tables which contain the coordinate

mapping for both coordinate directions in the image plane. The warping tables are calculated from disparity, which is accurately given due correspondence via the model, in straightforward fashion.

5

The main idea in **detection of obstacles** is to warp one image, e.g. the left image to the right one, and perform some comparison. The objects on the staircase, which should be detected, have the property of being closer to the camera than the staircase on which they are placed. Therefore, objects in the image being warped appear at different positions than they appear in the unwarped image. On the other hand, disturbances like dirt or inscriptions on the staircase, appear on the same position in warped and unwarped image.

10  
15

To summarise, an extension of stereo based obstacle detection procedures to regularly structured and non-flat back-ground was employed. **Grouping** based on a **cross ratio** constraint improved RANSAC sampling. **Pose estimation** provides externally calibrated cameras, which simplify and accelerate stereo processing and the object detection task which is performed using a **pyramid based segmentation procedure**. A high reliability of the approach was found experimentally, i.e. a rate of omission of an obstacle in the order of magnitude of 1 percent, and a rate of false detection of an obstacle in the order of magnitude of 5 percents. Cylindrical objects down to a size of less than 15 centimeters in height were detected reliably.

20  
25

30

The processing unit is connected through the control line 6 to the **escalator controller** 7 and can therefore control the restarting of the escalator after a stop in dependence on the detection of a person or obstacle on the escalator.

**Signal connections** between the PCs and the escalator control use simple wires, through which signals from the staircase control go to each PC and back. Signals from the PC to the control are combined in disjunctive fashion, e.g.  
5 an object is detected if any PC signals a detected object, etc.

Three output signals are provided to the control:

- 10 • Object detected.
- Warning, e.g. camera problem.
- Failure, i.e. system not working.

Additionally, the system should support a so-called test  
15 mode, where images are fed into the system from stored location and not from the cameras. Therefore, two input signals are necessary:

- Staircase in standstill.
- Test mode requested.

20 Signalling between monitoring and staircase control is done using the digital input/output channels of the NuDAO 6208 multi-channel analogue output card. Besides the analogue output channels, the NuDAO 6208 card provides four input and four output channels.

25

The controller is connected through the motor supply line 8 to the **escalator motor** 9 and can therefore restart the motor or keep it in a still position.

30 Further technical requirements for the monitoring system are that the system may consist of two independent channels or control units. A watchdog function is required, i.e. the system may be continuously checked for availability.

It is obvious for the man skilled in the art, that the disclosed system and method using pairs of stereoscopic images can be also used to detect persons and objects in an elevator car or in a lobby in front of elevator doors.

**References**

- 1 Escalator
- 2 Person
- 5 3.1, 3.2 Video camera pair
- 4 Linking means camera-processing unit
- 5 Processing unit
- 6 Control line
- 7 Escalator controller
- 10 8 Motor supply line
- 9 Escalator motor



**Claims**

1. Monitoring system for the detection of obstacles and persons comprising at least one video camera and escalators  
5 and/or moving walks,  
characterized in that,  
the monitoring system acquires stereoscopic images.

2. Monitoring system according to claim 1,  
10 characterized in that,  
the video cameras are located above the escalator and/or moving walk.

3. Monitoring system according to claim 1,  
15 characterized in that,  
the video cameras are located in the balustrade of the escalator and/or moving walk.

4. Monitoring system according to claims 1 to 3,  
20 characterized in that,  
more than one pair of video cameras are arranged along the escalator and/or moving walk in order to monitor conveniently the whole stroke of the escalator and/or moving walk.

25  
5. Monitoring system according to claims 1 to 4,  
characterized in that,  
the monitoring system comprises a processing unit processing the stereoscopic images.

30  
6. Monitoring system according to claims 1 to 5,  
characterized in that,  
the monitoring system comprises means for linking the video cameras with the processing unit, in particular a data

exchange bus, and/or means for storing the stereoscopic images.

7. Monitoring system according to claim 5 to 6,  
5 characterized in that,  
the processing unit is at least one personal computer loaded with an image processing software program basing on the principle to compare the digital data of a model image with those of an actual acquired image.

10

8. Monitoring system according to claim 5 to 6,  
characterized in that,  
the processing unit is integrated in the cameras.

15 9. Monitoring system according to claims 5 to 8,  
characterized in that,  
the monitoring system is connected electrically to an escalator and/or moving walk control, which restarts the escalator and/or moving walk after a stop only when no  
20 obstacle and/or person are detected on the escalator and/or moving walk.

10. Computer program product for the detection of obstacles and/or persons on escalators and/or moving walks,  
25 characterized in that,  
the computer program product loads in a processor and processes stereoscopic images of the escalator and/or moving walk.

30 11. Computer program product according to claim 10,  
characterized in that,  
the computer program product restarts the escalator and/or moving walk after a stop only when no obstacle and/or person are detected on the escalator and/or moving walk.



12. Method for the detection of obstacles and persons on escalators and/or moving walks, characterized in that,

- 5 at least one video camera acquires stereoscopic images and a processing unit processes these images.

13. Method according to claim 11, characterized in that,

- 10 the escalator and/or moving walk is restarted automatically after a stop only when no obstacle and/or person are detected on the escalator and/or moving walk.



**Abstract**

The invention discloses a monitoring system for the detection of obstacles and persons on escalators and/or  
5 moving walks comprising at least one video camera, whereby the monitoring system acquires stereoscopic images.

(Figure 1)



Fig. 1

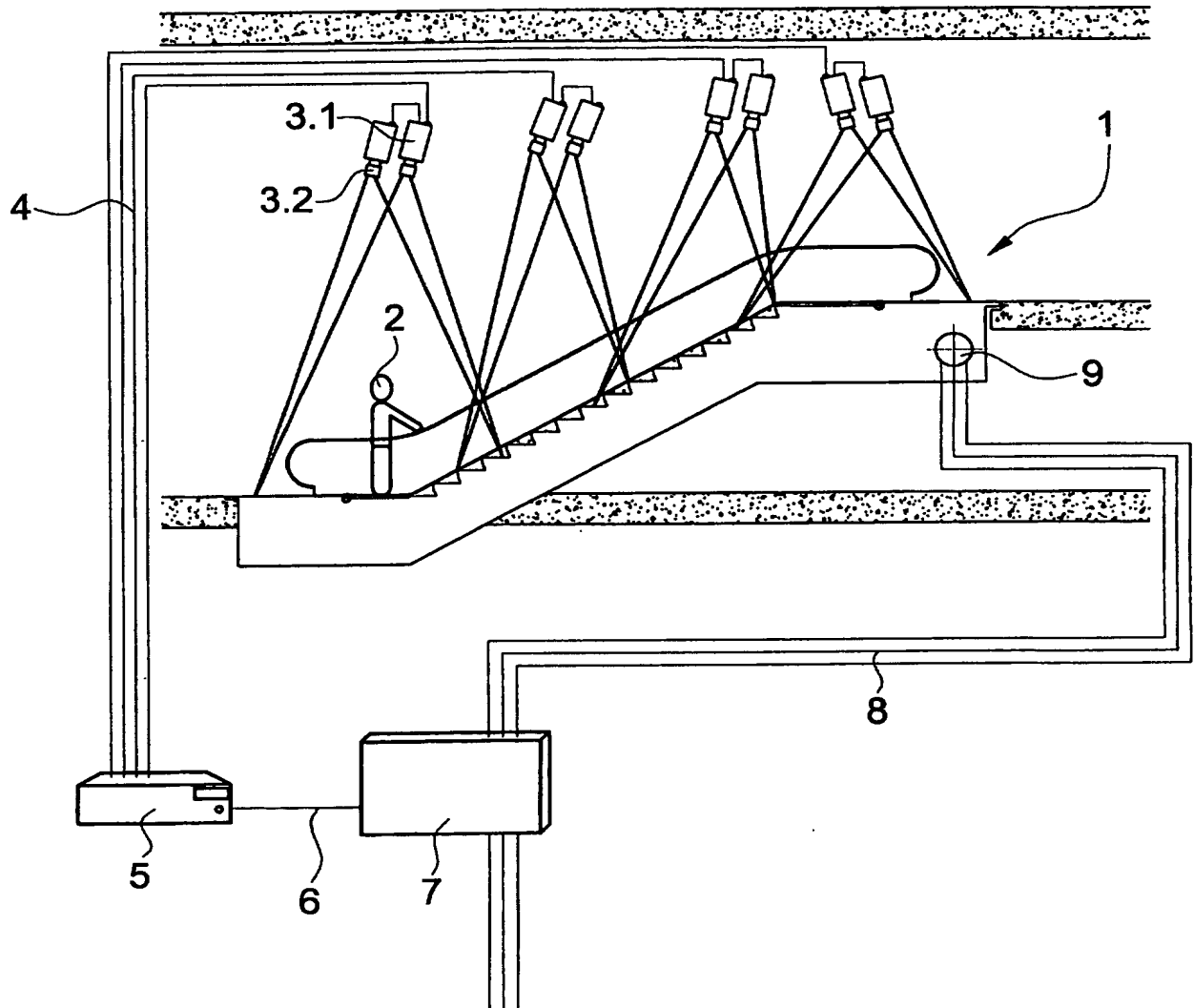


Fig. 2

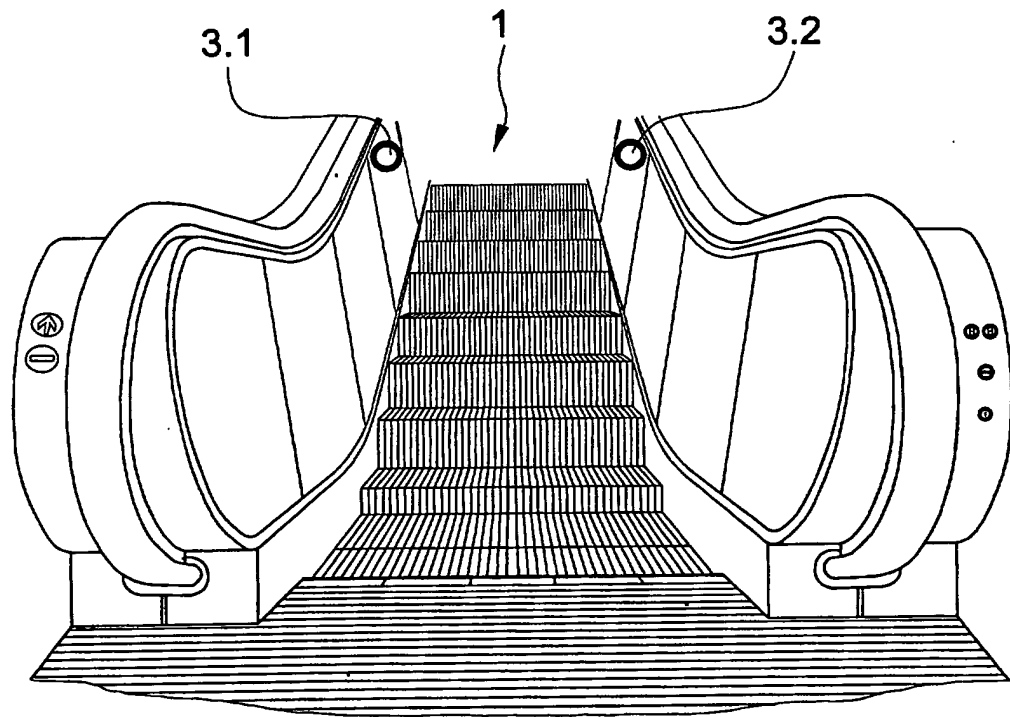


Fig. 3

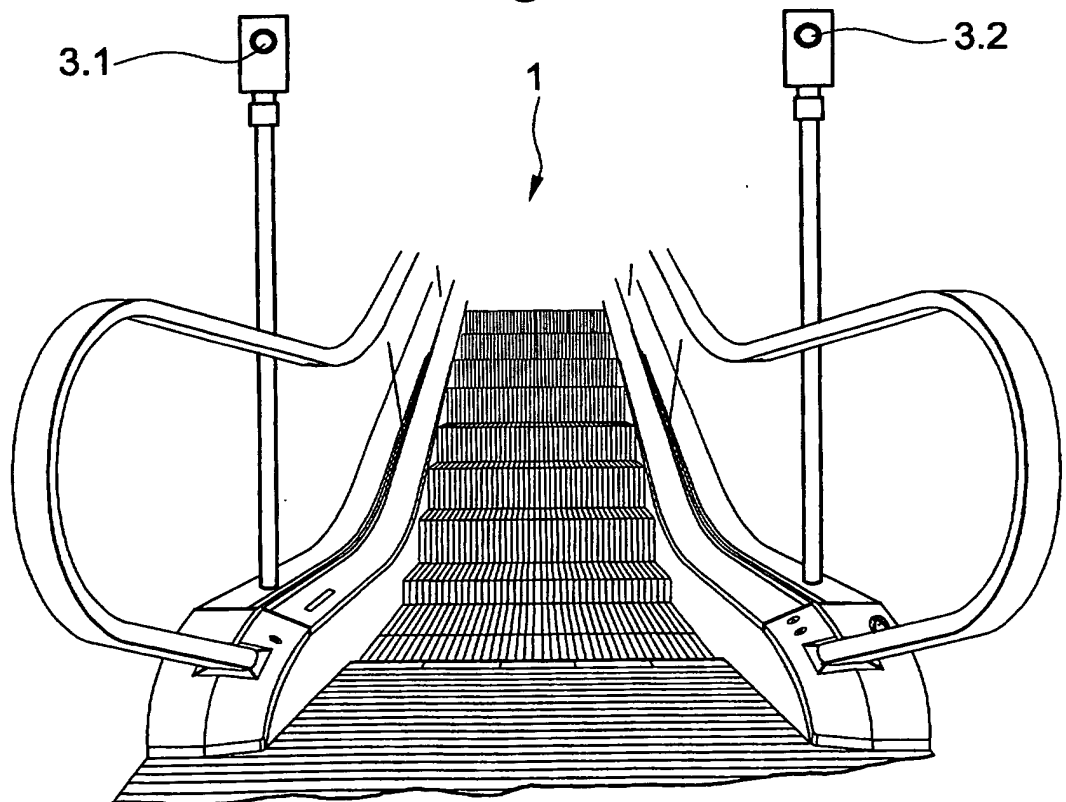


Fig. 4

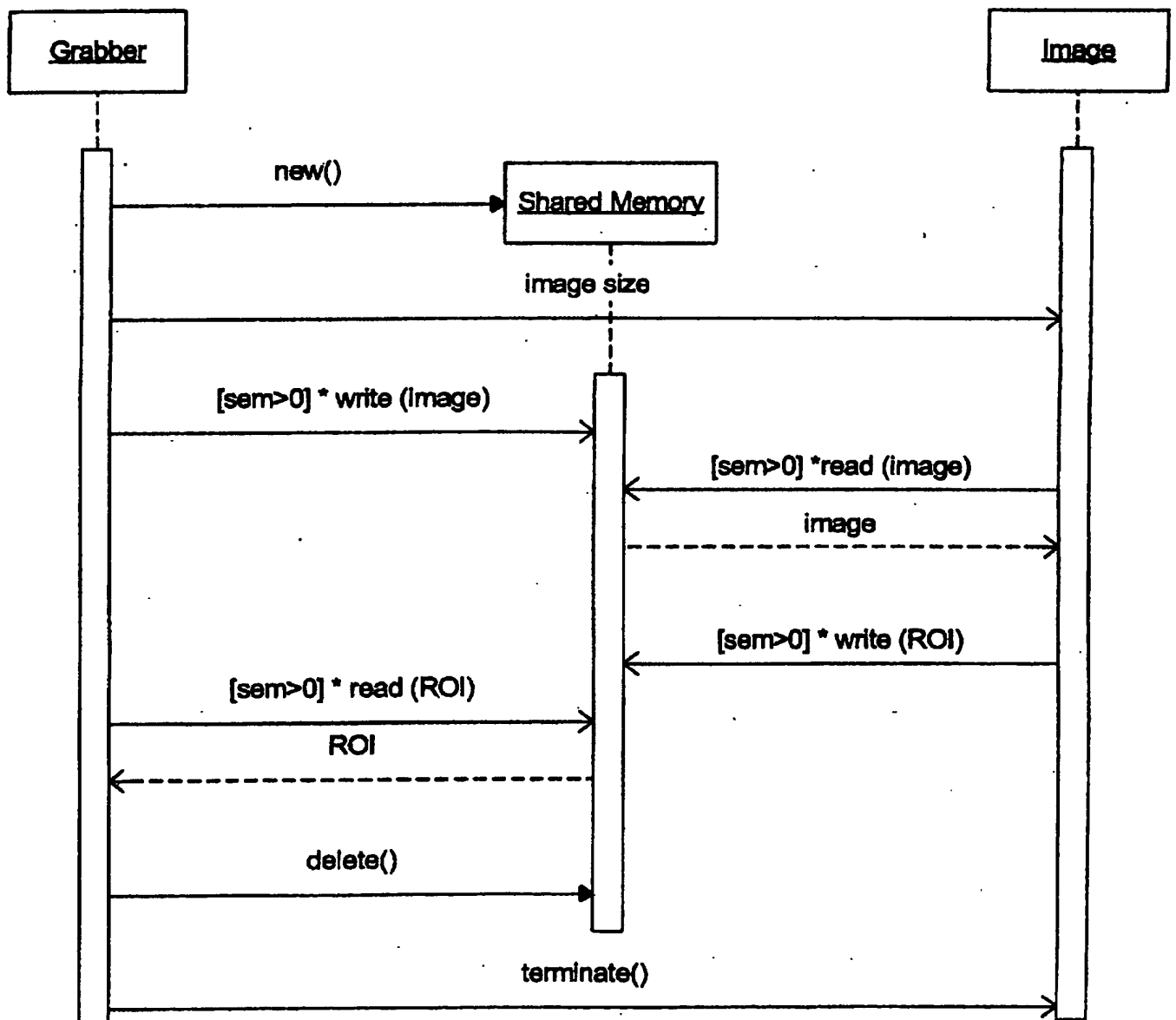


Fig. 5

